

ML Project Idea

Your Name

Motivation

Situation / Problem / Goal:

What do you want to accomplish, i.e.,
what problem do you want to solve /
what situation do you want to improve?

picture (optional)

Value Generation:

- process improvement (reduce costs)
- new product / feature / service (increase revenue)



Business KPI:


How can you measure/quantify the progress towards your goal?

Status Quo:





What is this KPI now?

Solution Outline

Deliverables:  **insights:** (generated by an one-off analysis of historical data)
 **software:** (to continuously make predictions for new data points)

 **Build**
 **Buy**

Inputs:

-  (numeric) values: _____
-  image
-  text
-  other: (e.g., audio, video)








1 Data Point:

What is one interaction that generates these measurements?



Workflow Integration:

Where does the input data come from? What happens to the ML outputs?

ML Solution & Output:

-  Dimensionality Reduction: **2D coordinates**
-  Outlier Detection: **anomaly score**
-  Clustering: **cluster index**
-  Regression: **continuous value:** _____
-  Classification: **discrete value (e.g., yes/no):** _____
-  Recommender Systems/Information Retrieval: **ranking of items**
-  Generative AI: (e.g., image, text, ...): _____

Additional Steps?

-  Explain predictions (e.g., to identify root causes)
-  Use model in optimization (to find optimal inputs)

Challenges & Risks (+ Mitigation Strategies)

What might go wrong?

Probability:

Can you do anything about this?

For example, not enough / not the right data
(=> How much data do you have??)

high/low

For example, improve the data infrastructure, systematically collect additional data, and automate data entry (incl. validation)

For example, what would be the worst case scenario when the model is wrong?

high/low

For example, add an option to keep a human in the loop when the model is unsure about its prediction

For example, legal/ethical concerns when working with user data

high/low

For example, external validation of the model to check for systematic bias and discrimination

For example, difficult to integrate the solution with existing systems, infrastructure, or processes

high/low

For example, invest in a platform team to centrally manage data integration and deployment and maintenance of ML models